From non-convexity to convexity: A game theoretic bridge

- In general, non-convex non-separable non-differentiable problems is known to be hard
- We start from separable smooth (continuously differentiable) non-convex functions with convex constraint
- Why non-convex is hard: Lagrangian-dual is used to convert a constrained problem to an unconstrained one, however non-convexity leads to duality gap; even unconstrained, multiple local extremes leads to sub-optimality
- Whether we can find a convex problem such that the optimizer of the non-convex one is consistent with the optimizer of the corresponding convex one?

Review of potential game theory

- A Game $G =< \mathrm{M}, \{u_m\}_{m\in M}, \{\Pi_m\}_{m\in M} >$ (M is the user set, u_m is the local objective for user m, and Π_m is the constraint set for user m) is a potential game if there is a function $\Phi : \Pi = \otimes \Pi_m \rightarrow \Re$ satisfying

$$\Phi(\vec{x}_m, \vec{x}_{-m}) - \Phi(\vec{y}_m, \vec{x}_{-m}) = u_m(\vec{x}_m, \vec{x}_{-m}) - u_m(\vec{y}_m, \vec{x}_{-m}), \forall m \in M, x_m, y_m \in \Pi_m, \vec{x}_{-m} \in \Pi_{-m}$$

  i.e., $\frac{\partial}{\partial x_m}\Phi(\vec{x}) = \frac{\partial}{x_m} u_m(\vec{x}_m), \forall m$ when Φ is smooth. Φ is the corresponding potential function.
- When Φ is concave ad smooth, and Π is convex, the set of maximums of Φ is consistent with the set of equilibriums of G

First try

- If we can construct a potential game with a strictly concave potential, and the global optimizer of the non-concave problem is an equilibrium of the potential game. Then this global optimizer is consistent with the maximum of the concave potential function
- If we let the global objective be the local objective for all users, then the global optimizer is an equilibrium of this game. However, this game does not have concave potential
- Operations (e.g., change of variables) that preserve the nature of the problem can not convert a con-convex problem to convex. Need to introduce auxiliary variable

A class of non-convex functions

- Optimization Problem: $\max_{\vec{x} \in \prod_x} f(\vec{x})$
- Variable $\vec{x} = (x_1, x_2, \ldots, x_L)$ is an L dimensional vector
- Constraint $\prod_x$ is convex
- The objective $f(\vec{x})$ is bridgeable over the constraint if
  - Continuously differentiable (smooth)
  - Positive inherently separable
  - Seemingly strictly concave

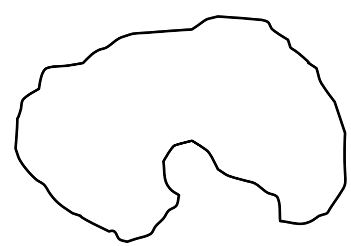
Inherently separable

- A function $f(\vec{x})$ is inherently separable if it can be written as

$$f(\vec{x}) = \sum_{l=1}^{L} f_l(K_l - \varphi_l(K_l)) \qquad \Pi_x \underset{[A]^{-1}}{\overset{[A]}{\rightleftarrows}} \Pi_K$$

  - $\vec{K} = [A]\vec{x}$, where $\det[A] \neq 0$, i.e., $\vec{x} \leftrightarrow \vec{K}$ is linear injection . Thus, Π_K, constraint of $\vec{K}$, is also convex, and $\vec{K}$ can be viewed as the new basis
  - φ_l are arbitrary differentiable functions. We call φ_l "separating functions"
  - Separable ->inherently separable

$$\Pi_K \underset{\varphi^{-1}}{\overset{\varphi}{\rightleftarrows}}$$



Positive inherently separable

- A function $f(\vec{x})$ is positive inherently separable if it is inherently separable and the separating functions satisfy (strictly positive inherently separable if the inequality holds strictly):

$$0 \le \frac{d}{dK_l}\varphi_l(K_l) \le 1, \forall l, \forall K_l \in [c^K_{l,m}, c^K_{l,M}]$$

- Let $Q_l = \varphi_l(K_l)$, then $K_l = \varphi_l^{-1}(Q_l)$. By applying derivative of inverse function, $\frac{d}{dQ_l}\varphi_l^{-1}(Q_l) = \frac{1}{\varphi_l'[\varphi_l^{-1}(Q_l)]}$ the above condition is equiv to

$$\frac{d}{dQ_l}\varphi_l^{-1}(Q_l) \ge 1, \forall l, \forall Q_l \in [c^Q_{l,m}, c^Q_{l,M}]$$

Seemingly concave

- If $f(\vec{x})$ is positive inherently separable, i.e.,

$$f(\vec{x}) = \sum_{l=1}^{L} f_l(K_l - \varphi_l(K_l)),\ \vec{K} = [A]\vec{x},\ \det[A] \neq 0;$$

$$0 \leq \frac{d}{dK_l}\varphi_l(K_l) \leq 1, \forall l, \forall \vec{K} \in \Pi_K,$$

and if $f_l(\bullet), \forall l$ are (strictly) concave, then $f(\vec{x})$ is seemingly (strictly) concave

Examples

- $(x_1+2x_2-e^{x_1+2x_2})^2+(x_1+x_2-e^{x_1+x_2})^2, \vec{x}\in\Re$ is inherently separable but not positive inherently separable
- $\log[1+x_1-\log(1+x_1)]+\log[1+x_2-\log(1+x_2)], \vec{x}\in\Re^+$ is separable, then inherently separable; and it is also strictly positive inherently separable, and seemingly strictly concave
- $f(x)=(1+e^{-x})^{-1}, x\in\Re$ (sigmoid) is positive inherently separable and seemingly strictly concave:

$(1+e^{-x})^{-1}=\log\{K-[K-e^{(1+e^{-K/M})^{-1}}]\}, K=Mx,\ \varphi(K)=K-e^{f}$

$\varphi'=1-\frac{1}{M}f'e^{f}$. This function, f, f' are bounded and $f'\geq 0$, so it is possible to find M

Constructing potential game

- Rewrite $f(\vec{x}) = \sum_{l=1}^{L} f_l(K_l - \varphi_l(K_l)),$
- Game

$$\max_{\vec{K} \in \Pi_K} \sum_{l=1}^{L} f_l(K_l - Q_l) = u_{\vec{K}}$$

$$\max_{\vec{Q}} \sum_{l=1}^{L} [f_l(K_l - Q_l) + g_l(Q_l)] = u_{\vec{Q}}$$

$\vec{Q}$ *are auxiliary variables*

Constructing potential game (Cont.)

- Auxiliary variable $\vec{Q}$ is unconstrained, necessary condition

$$\frac{\partial}{\partial Q_l}\sum_{l=1}^{L}[f_l(K_l - Q_l) + g_l(Q_l)] = 0, \forall l$$

and require $Q_l = \varphi_l(K_l)$ to always be the solution

- This means

$$\frac{\partial}{\partial Q_l}\sum_{l=1}^{L}[f_l(K_l - Q_l) + g_l(Q_l)]$$

$$= \frac{\partial}{\partial Q_l}[f_l(K_l - Q_l) + g_l(Q_l)]$$

require strictly concavity of f so that f' is not constant

$$= -f_l'[\varphi^{-1}(Q_l) - Q_l] + g_l'(Q_l) \equiv 0$$

Constructing potential game (Cont.)

- Thus, for $\forall Q_l \in [c_{l,m}^{Q}, c_{l,M}^{Q}]$,

$$g_l''(Q_l) = [\frac{d}{dQ_l}\varphi^{-1}(Q_l) - 1] f_l''[\varphi^{-1}(Q_l) - Q_l] \leq 0$$

f is strictly concave, positive inherently separable leads to concavity of g, and strictly positive inherently separable leads to strictly concavity of g

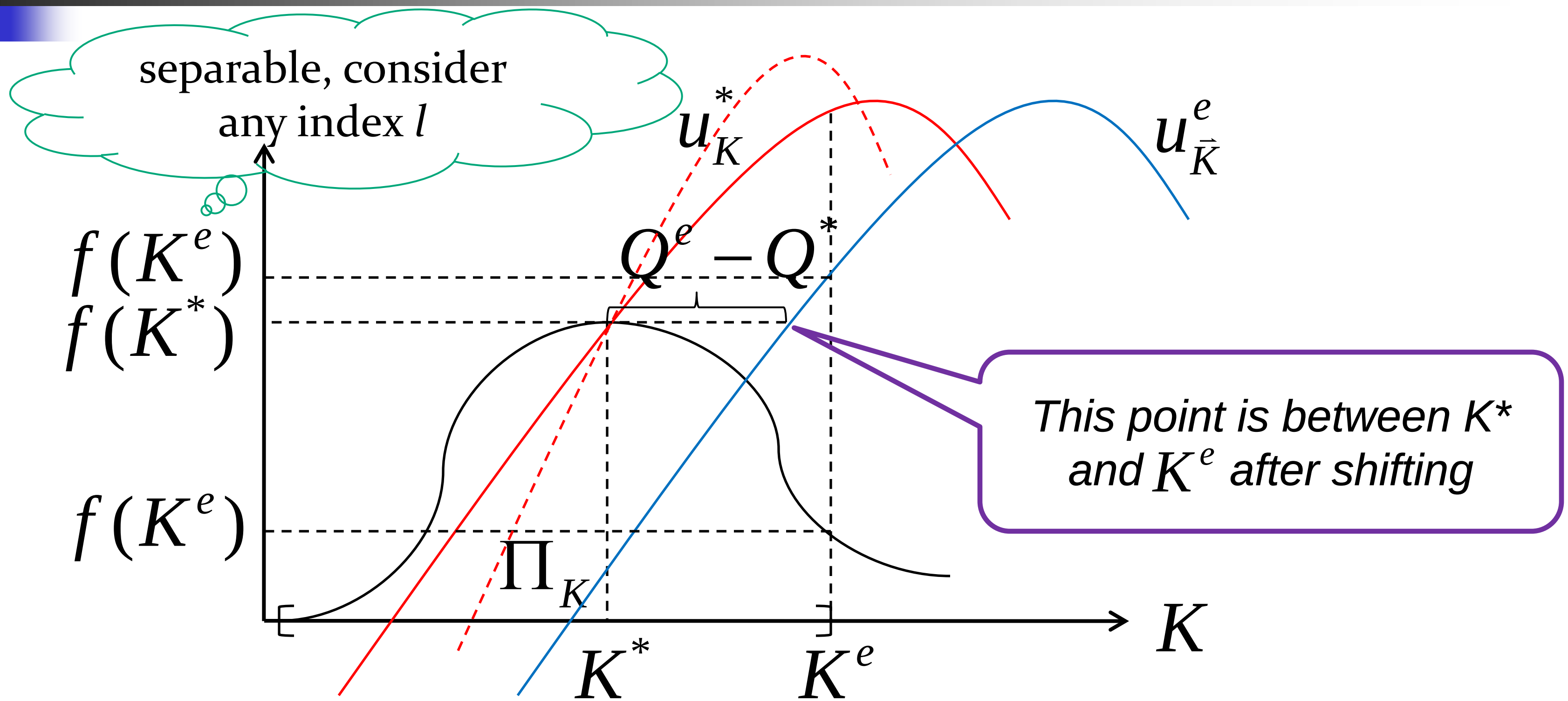
- Lemma: the global optimizer of bridgeable functions over a convex set is an equilibrium of the constructed potential game

Proof outline of the lemma

- Given optimizer $\vec{x}^*$, let $(\vec{K}^*, \vec{Q}^*)$ denote the corresponding point in the game. Suppose it is not an equilibrium, then $\vec{K}^*$ is not the optimizer of $u_{\vec{K}}^* = \sum_{l=1}^{L} f_l(K_l - Q_l^*)$ under constraint $\prod_{\vec{K}}$
- Let $(\vec{K}^e, \vec{Q}^e)$ be the equilibrium, this means $\vec{K}^e$ is the optimizer of $u_{\vec{K}}^e = \sum_{l=1}^{L} f_l(K_l - Q_l^e)$ under $\prod_{\vec{K}}$
- Note

$$\sum_{l=1}^{L} f_l(K_l^* - Q_l^*) = f(\vec{x}^*),\ \sum_{l=1}^{L} f_l(K_l^e - Q_l^e) = f(\vec{x}^e)$$

Proof outline of the lemma (Cont.)



- Note $0 \le \frac{dQ}{dK} = \frac{d}{dK}\varphi(K) \le 1$, we will reach a contradiction $f(K^*) < f(K^e)$

Bridge from non-convex to convex

- Theorem: If $f(\vec{x})$ is bridgeable, by writing $f(\vec{x})$ as $f(\vec{x}) = \sum_{l=1}^{L} f_l(K_l - \varphi_l(K_l)), \vec{K} = [A]\vec{x}$, and letting $\Phi(\vec{K},\vec{Q}) = \sum_{l=1}^{L} [f_l(K_l - Q_l) + g_l(Q_l)]$, where $g_l'(Q_l) \equiv f_l'[\varphi^{-1}(Q_l) - Q_l]$, then $\Phi(\vec{K},\vec{Q})$ is concave. Let $\{\vec{K}^m,\vec{Q}^m\}$ be the set of optimizer for $\max_{\vec{K}\in\Pi_K,\vec{Q}\in\otimes_l[c_{l,m}^Q,c_{l,M}^Q]} \Phi(\vec{K},\vec{Q})$ and let $M = \{\vec{x}^m : \vec{x}^m = [A]^{-1}\vec{K}^m\}$. Then, the global optimizer of $\max_{\vec{x}\in\Pi_x} f(\vec{x})$ where Π_x convex, belongs to M.

Bridge from non-convex to convex (Cont.)

- The theorem can be proved using the previous arguments and lemma, combined with the potential game theory
- Corollary: If $f(\vec{x})$ is strictly positive inherently separable, i.e., the separating functions satisfy $0 \le \varphi_l' < 1, \forall l$, then $\Phi(\vec{K},\vec{Q})$ is strictly concave, and the global optimizer of $f(\vec{x})$ is mapped to the unique optimizer of $\Phi(\vec{K},\vec{Q})$
- The strictly concavity can be verified from the following Hessian

$$\begin{bmatrix} f_l''(K_l - Q_l) & -f_l''(K_l - Q_l) \\ -f_l''(K_l - Q_l) & f_l''(K_l - Q_l) + g_l''(Q_l) \end{bmatrix} \to \begin{bmatrix} f_l''(K_l - Q_l) & 0 \\ 0 & g_l''(Q_l) \end{bmatrix}$$

Handling non-strictly concavity

- If $\Phi(\vec{K},\vec{Q})$ is not strictly concave, from the previous discussion, this non-strictly concavity is due to $g_l(Q_l)$, by adding quadratic terms $-\sum_{l=1}^{L}(\tilde{Q}_l - Q_l)^2$, we can handle it by using proximal algorithm
- The bridging method is not operations that preserves the nature of an optimization problem. It introduces auxiliary variables and the potential function is not an transformation from the original objective
- Obtaining the explicit expression of $g_l(Q_l)$ from $g_l'(Q_l) = f_l'[\varphi^{-1}(Q_l) - Q_l]$ and $\varphi_l^{-1}(\bullet)$ can sometimes be difficult, do we need to get $g_l(Q_l)$ and $\varphi_l^{-1}(\bullet)$ in the algorithm? Fortunately, no

Implementation

- Without loss of generality, let $\Pi_K = \{\vec{K} : [C]\vec{K} \le \vec{c}_M\}$ which satisfies the Slater's condition, then the Lagrangian
$$L(\vec{K},\vec{Q},\vec{\lambda}) = \Phi(\vec{K},\vec{Q}) - \vec{\lambda}^T([C]\vec{K} - \vec{c}_M)$$
- In the dual algorithm, at each iteration m, given $\vec{\lambda}^{(m)} \ge 0$, need to solve
$$\max_{\vec{K},\vec{Q}}\{\Phi(\vec{K},\vec{Q}) - \vec{\lambda}^T[C]\vec{K}\}$$
- $$\frac{\partial}{\partial Q_l}\{\Phi(\vec{K},\vec{Q}) - \vec{\lambda}^T[C]\vec{K}\} = -f_l'[K_l - Q_l] + g_l'(Q_l) = 0$$
$$\Rightarrow Q_l = \varphi_l(K_l)$$

Implementation (Cont.)

characterize the duality gap of the original non-convex problem

- $$\frac{\partial}{\partial K_l}\{\sum_{l=1}^{L}[f_l(K_l-\varphi_l(K_l))+g_l(\varphi_l(K_l))]-\vec{\lambda}^T[C]\vec{K}\}$$
$$=(1-\varphi_l'(K_l))f_l'(K_l-\varphi_l(K_l))+\varphi_l'(K_l)g_l'(\varphi_l(K_l))-\sum_{k=1}^{L}\lambda_k C_{kl}$$
$$=f_l'(K_l-\varphi_l(K_l))-\sum_{k=1}^{L}\lambda_k C_{kl}=0$$
- $K_l^{(m),temp}$ can be calculated using Newton's method without knowing the expression of $g_l(\bullet)$ and $\varphi_l^{-1}(\bullet)$
- Even when K_l has extra simple constraint $c_{l,m}\le K_l\le c_{l,M}$, due to concavity:

$$K_l^{(m)}=\begin{cases} c_{l,m}, & \text{if } K_l^{(m),temp}\le c_{l,m} \\ c_{l,M}, & \text{if } K_l^{(m),temp}\ge c_{l,m} \\ K_l^{(m),temp}, & \text{o/w}\end{cases}$$

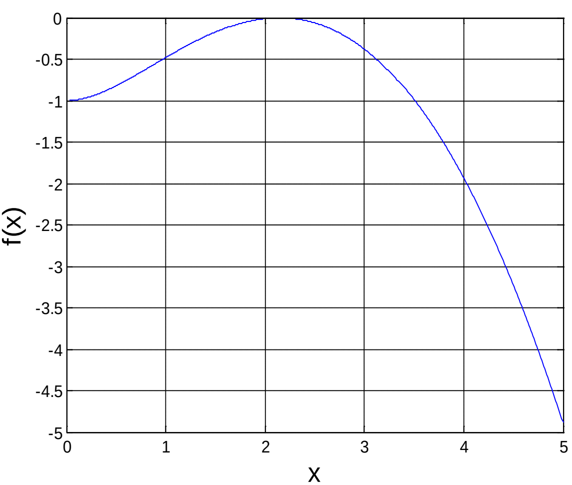
Illustration of bridging

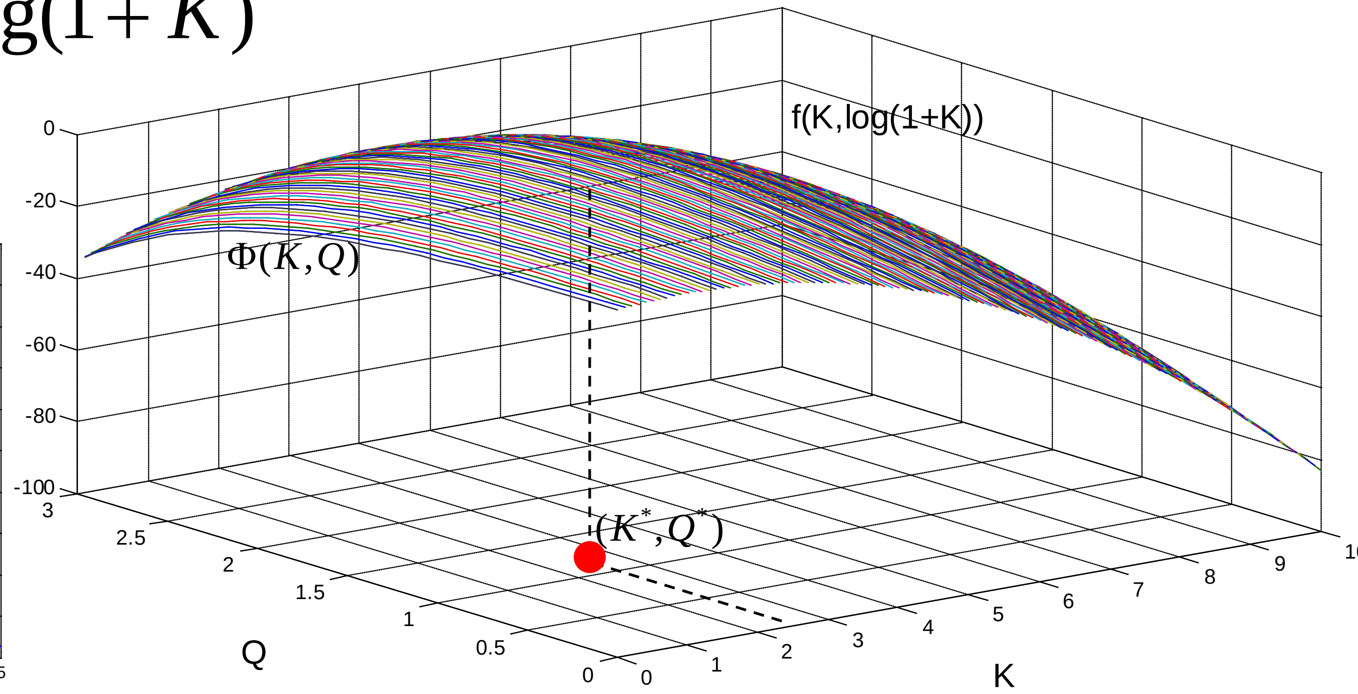
$$f(x) = -[x - \log(1+x) - 1]^2, x \geq 0 \qquad x^* = 2.1462$$

$$\Phi(K,Q) = -[K - Q - 1]^2 + Q^2 + 4Q - 2e^Q \qquad K^m = 2.146$$

$$Q = \varphi(K) = \log(1+K)$$

$$K = x \geq 0$$





How large the class of functions is

- Recall that the class of functions we are interested (define as bridgeable functions) in has the following properties:
  - Continuously differentiable (smooth)
  - Positive inherently separable
  - Seemingly strictly concave
- It is apparent that inherently separable concave functions are included
- From previous examples, we observe that some sigmoid functions belong to this class, the first question is whether the whole sigmoid class is contained?

Monotone and Sigmoid functions

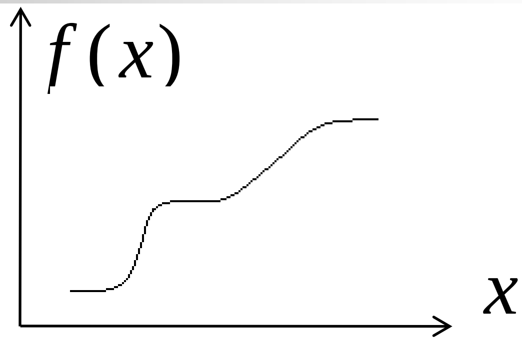
- Some properties of sigmoid functions $f(x)$:
  - "S" or opposite "S" shaped: either $f'(x) \geq 0$ or $f'(x) \leq 0$
  - Bounded: $|f(x)| \leq c_1 < \infty$ and $|f'(x)| \leq c_2 < \infty$
  - Have one inflection point
- Proposition: The class of inherently separable monotone functions with bounded function value and bounded first derivative belongs to the class of bridgeable functions. Strictly monotone leads to strictly positive inherently separable.
- Corollary: The class of inherently separable sigmoid functions belongs to the class of bridgeable functions

Proof outline of the proposition

- Note $f(x) = \log\{x - [x - e^{f(x)}]\} = \log\{-x - [-x - e^{f(x)}]\}$
  - If $f'(x) \geq 0$, let $K = Mx, M > 0$, then $f(K) = \log\{K - [K - e^{f(K/M)}]\}, \varphi(K) = K - e^{f(K/M)}$ $\varphi' = 1 - f'e^{f} / M$. Choose M so that $\varphi' \in [0,1]$
  - If $f'(x) \leq 0$, let $K = -Mx, M > 0$, then $f(K) = \log\{K - [K - e^{f(-K/M)}]\}, \varphi(K) = K - e^{f(-K/M)}$ $\varphi' = 1 + f'e^{f} / M \leq 1$. Choose M so that $\varphi' \in [0,1]$
  - Sigmoid function optimization under convex constraints is a non-convex problem in nature. This tells us that the game theoretic bridging is unlike change of variables etc, that preserves the nature of the function

More than monotone and sigmoid functions

- Allow multiple inflection points:



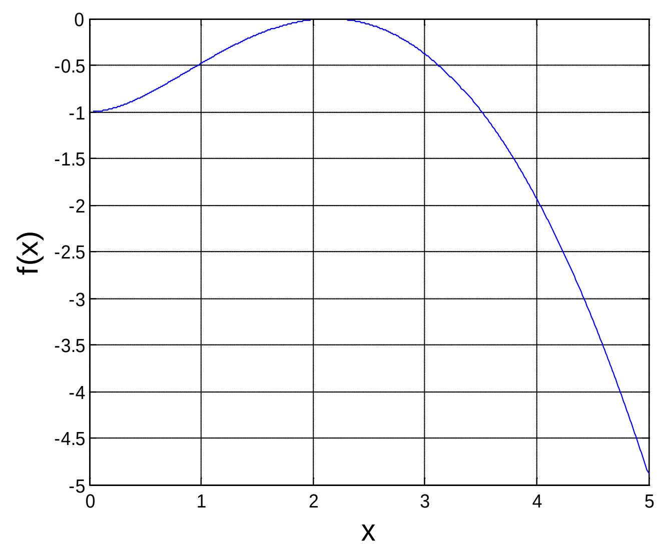
- Allow unbounded value and unbounded first derivative , do not need to be monotone:

$$f(x) = -[x - \log(1+x) - 1]^2, x \geq 0$$

$$Q = \varphi(K) = \log(1+K)$$

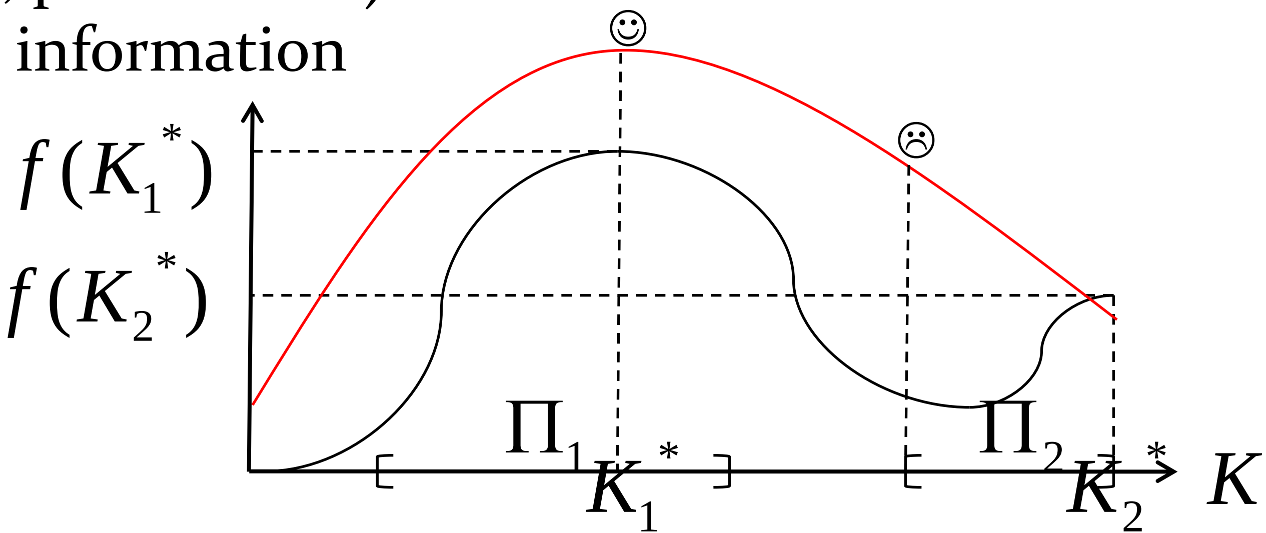
$$K = x \geq 0$$

$$f(\bullet) = -[\bullet - 1]^2$$



An outer bound for bridgeable functions

- Theorem: A bridgeable function has at most one local maxima
- Proof outline: all bridgeable function optimization under convex constraint can be bridged to a concave potential function optimization under the corresponding convex constraint; potential objective construction does not use constraint information



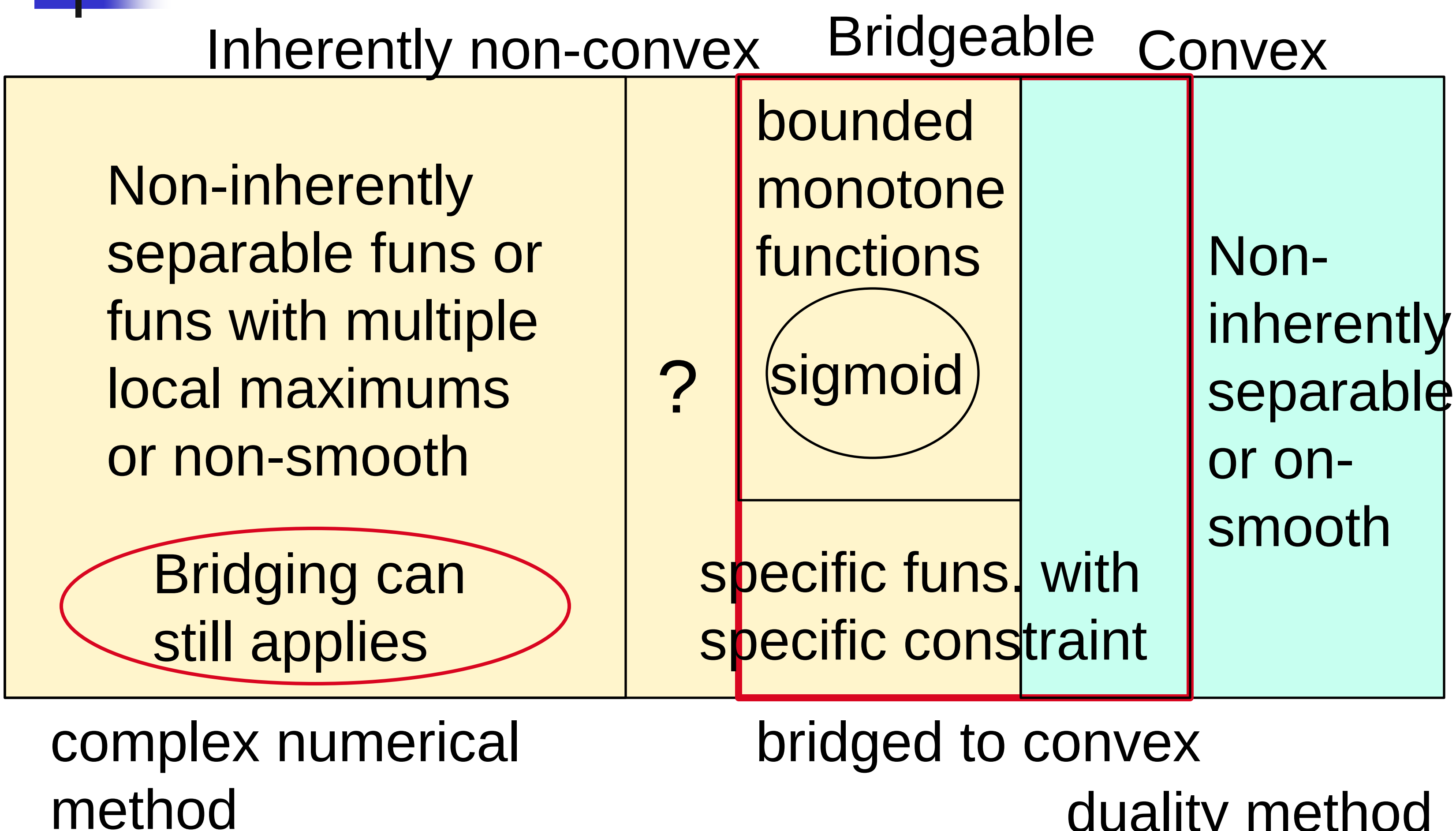
Unknown gap

- Example: $f(x) = \exp[\log(x) - x], x \geq 0$

  $$f' \geq 0, x \in [0,1];\ f' \leq 0, x \in [1,+\infty]$$

  This function has only one local maxima, however, it is not bridgeable over the whole constraint
- Bridgeable or not is highly related to the constraint set, so it should be evaluated under the constraint. Recall that $\bigotimes_l [c^K_{l,m}, c^K_{l,M}]$ is the region in which $0 \leq \frac{d}{dK_l}\varphi_l(K_l) \leq 1, \forall l$. Thus, if $\Pi_K \subseteq \bigotimes_l [c^K_{l,m}, c^K_{l,M}] \neq \Phi$, then $f(x)$ is bridgeable over Π_x, otherwise, it is only bridgeable over a subset mapped from $\bigotimes_l [c^K_{l,m}, c^K_{l,M}]$

Illustration of the bridgeable class



Application: Network utility maximization (NUM)

- Problem formulation: $$\max_{\vec{x}\in\Pi} \sum_{l=1}^{L} U_l(x_l)$$

  where Π is a compact convex set, and $U_l(\bullet)$ are bridgeable. Physically, x_l can be link rate, net gain, cost etc., the constraint can be capacity constraint and other QoS constraint etc. All NUM type formulations have separable objectives
- This problem is ready to be solved in a *distributed* way with the previous preparation
- Except for the convergence gap due to the dual algorithm, there is no theoretical gap of the solution

Comparison with existing method

- [Lee, Mazumdar, Shroff] 05, Non-Convex Optimization and Rate Control for Multi-Class Services in the Internet : distributed, but asymptotically optimal with increasing number of users, and only deal with sigmoid
- [Fazel, Chiang] 06, Network Utility Maximization With Non-concave Utilities Using Sum-of-Squares Method: consider polynomial (possibly multiple extremes) objective, but using convex relaxation methods (only achieves an upper bound, i.e., there is a theoretical gap), and centralized

Can we do more: non-separable but convex

- Bridging method can be applied to optimizations of bridgeable functions under convex constraint. Whether this method can be used for other functions?
- All convex problems, no matter separable or not, can be viewed as a special case which is already bridged, i.e., no auxiliary variables
- In most practical scenarios, we are interested in a compact set and functions with bounded derivative, so the boundedness is not a big issue
- Difficulties: multiple local extremes and non-separable

Can we do more: non-separable non-convex

- A general version of the bridging theorem: If $f(\vec{x})$ can be written as $F(\vec{K}-\vec{Q})$, where the injection $\Pi_x \rightarrow \Pi_K$ preserve convexity , F is continuously differentiable and concave, $\frac{\partial}{\partial Q_l}F$ are not constant, and $\vec{Q}=\vec{\varphi}(\vec{K})$ is an function satisfying

$$\gamma(\vec{K}_1-\vec{K}_2)=\vec{\varphi}(\vec{K}_1)-\vec{\varphi}(\vec{K}_2), \forall \vec{K}_1, \vec{K}_1 \in \Pi_K \quad (1)$$

where $0 \le \gamma \le 1$ is some constant. Further, if there is a concave function $g(\vec{Q})$ such that

$$\frac{\partial}{\partial Q_l}F(\vec{\varphi}^{-1}(\vec{Q})-\vec{Q})+\frac{\partial}{\partial Q_l}g(\vec{Q})\equiv 0, \forall l \quad (2)$$

then maximizing $f(\vec{x})$ over convex Π_x can be bridged to a convex problem

In general, it's extremely difficult to find a construction that satisfies both (1) & (2)!

Can we do more: separable with multiple extremes

- Example: $\max_{0.5 \le x \ge 1.5} f(x)$, where $f(x) = x - \log(x)$

  $f' \le 0, x \in [0.5,1]; f' \ge 0, x \in [1,1.5]$

- Let $K = -Mx, \forall x \in [0,1]; K = Mx, \forall x \in [1,+\infty]$, then for each set, the function is bridgeable. Solve for both sets and get the maximum
- The number of sets for L dimensional $\vec{x}$ is 2^L (exponential)!

Can we do more: separable with multiple extremes (Cont.)

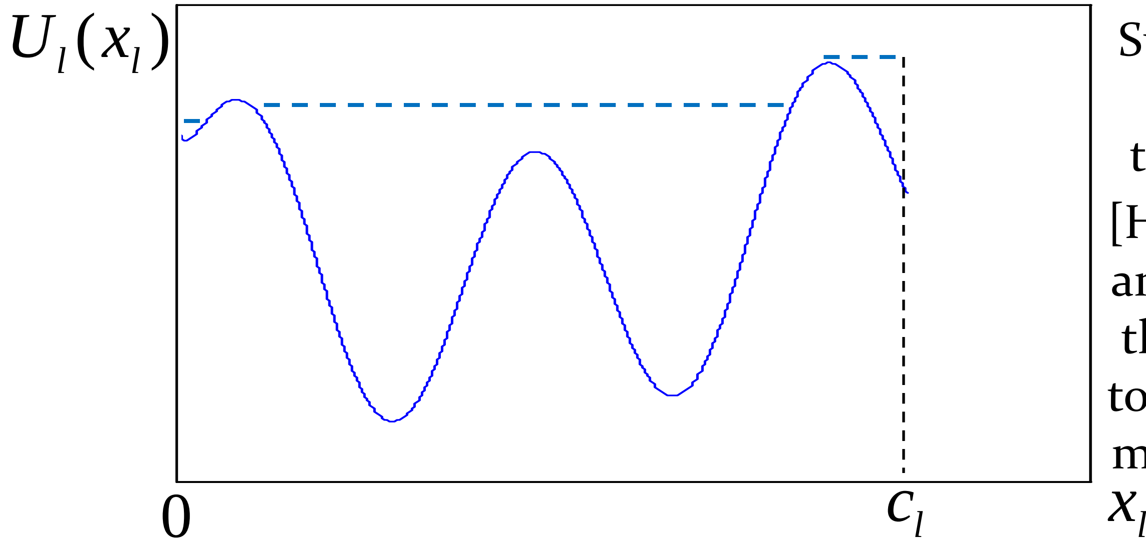
- The conditions in the general bridging theorem are difficult to verify, however, when the function is inherently separable, we only need to check the first derivative of each separating function φ_l. From the previous example, we can see that multiple extremes leads to unbridgeable due to not being monotone .
- Without loss of generality, we consider a NUM

$$\max_{\vec{x}} \sum_{l=1}^{L} U_l(x_l) \quad \text{s.t. } 0 \le x_l \le c_l, [H]\vec{x} \le \vec{C}$$

where $[H]$ has nonnegative entries, e.g., routing matrix and $\vec{C}$ is component-wisely nonnegative, e.g., capacity

Future work: Can we do more: separable with multiple extremes (Cont.)

- $U_l(x_l)$: bounded value and first derivative
- Rough idea: from non-bridgeable to bridgeable



Starting from zero, construct an increasing function $\tilde{U}_l$ that covers the original one; [H] has nonnegative entries, any maximum will not lie in the dotted area when using $\tilde{U}_l$ to run the optimization, this means we can optimize over $\tilde{U}_l$

- We construct increasing ($K_l = Mx_l, M > 0,\ [\tilde{H}]$ still has nonnegative entries, and $\varphi_l(K_l) = K_l - e^{\tilde{U}_l(K_l/M)}$) functions, but it is not differentiable

Future work: Can we do more: separable with multiple extremes (Cont.)

- Rough idea of handling non differentiability: Let N_l be the number of non-differentiable points of $\tilde{U}_l$, then there are $\prod_l N_l$ non differentiable points of the sum objective and $\prod_l (1+N_l)$ differentiable regions. At each region, do the bridging and run dual method. Recall that for the primal maximization over K_l , by setting the first derivative zero, we get $f_l'(K_l - \varphi_l(K_l)) - \sum_{k=1}^{L} \lambda_k \tilde{H}_{kl} = 0$. Note that $f_l'(\bullet) = \frac{1}{\bullet}$ and $\varphi_l(K_l)$ is continuous over the whole constraint. This means the points around the non-differentiable points have consistent primal update equation. It is natural to expect that the potential function will be continuous and concave with $\prod_l N_l$ non-differentiable pts, and sub-gradient method may apply

Application: Downlink power allocation under SINR model

- Same model as in [Lee, Mazumdar, Shroff] 05 , Downlink Power Allocation for Multi-class Wireless Systems

$$\max_{\vec{P}} \sum_{i=1}^{N} U_i(\gamma_i(\vec{P})), \text{ s.t. } 0 \le P_i \le P_T, \sum_i P_i \le P_T; \;\; \gamma_i(\vec{P}) = \frac{N_i P_i}{\theta(\sum_m P_m - P_i) + A_i}$$

- This is not separable, however, it is shown in the above paper that at the optimum, $\sum_i P_i = P_T$
- Then under $\sum_i P_i = P_T$, as long as the objective is bridgeable, we are done
- In general non-separable non-convex is extremely difficult, for specific problem, we can make non-separable problem into a separable format

Application: Downlink power allocation under SINR model (Cont.)

- To be continue: how about there is linear cost term in the utility, then it is not true the optimum is obtained under $\sum_i P_i = P_T$
- Rough idea
  - Fix total power
  - Add back the term that captures the duality gap
  - Search through the total power-one dimension

Future work: Can we do more: from non-separable to separable

- General non-separable power allocation formulation

$$\max_{\vec{P}} \sum_{l} \log\left(1+\frac{P_l}{\sigma_l+\sum_{i\neq l} h_{il}P_i}\right)-C_l P_l, \text{ s.t. } [A]\vec{P}\leq \vec{B}, P_{l,m}\leq P_l \leq P_{l,M}$$

where $[A], \vec{B}, \vec{C}, [H], \vec{\sigma}$ are all nonnegative

- To be continued…

Q&A

Thank you !